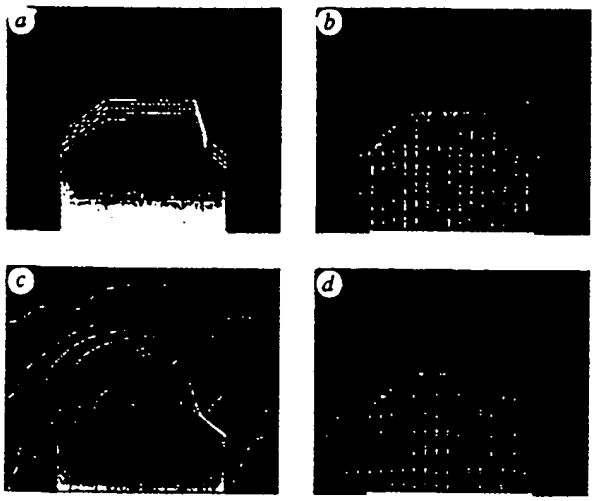

FIG. 4 Plan view of input buoyancy structure and faults (a), model output of surface velocity (red arrows) and μ_{eff} (b), topography and long-wavelength (>1,000 km) geoid (c) for the model with variable strikes of slabs, and surface velocity and μ_{eff} (d) for the model excluding the faults. Contour interval for geoid is 10 m and dashed lines are negative values. The colour scales are the same as in Fig. 2. The length, width and thickness of the model boxes (in km) are 6,000, 7,500 and 1,500, respectively, which are covered by 54 × 54 × 16 elements.

faults and slabs with variable strikes (Fig. 4a) shows that surface velocities remain plate-like (Fig. 4b) with $P = 0.87$ and $R_{T;P} = 0.40$ ($R_{T/P} = 0.42$ for a perfect plate). Surface velocities in R are predominantly parallel to the vertical faults and show oblique subduction where the dipping faults are not parallel to the spreading centre (Fig. 4b). Trench topography is evident oceanward from the dipping faults and there is 40 m long-wavelength geoid high over slabs (Fig. 4c). Qualitatively, both trench/fore-bulge topography and the geoid are essentially identical to observed features in the western Pacific (for example, approximately 4 km deep and 150 km wide trenches, and 30 m long-wavelength geoid high). The reduced P results form the velocity variations in the spreading perpendicular direction near the short vertical fault (Fig. 4b) where slabs are younger; a more realistic temperature-dependent rheology may lead to a higher plateness. When the faults are excluded from the model, surface velocities vary gradually in both the spreading parallel and perpendicular directions with P decreased to 0.75 (Fig. 4d). However, $R_{T/P}$ only decreases slightly to 0.39, suggesting that a large fraction of toroidal power is not a sufficient criterion for characterizing the existence of plates.

These models show that the characteristics of plate motion can be explained with realistic mantle buoyancy, weak faults and a power-law rheology ($n = 3$). Besides plate kinematics, the models also simultaneously predict realistic trench/fore-bulge topography and the long-wavelength geoid (Fig. 4b and c). In the past, different physical processes were used to explain these observed features (for example, elastic plate bending models for trench/fore-bulge topography and laterally homogenous viscous flow models for the geoid). By simultaneously matching observed trench topography, plate kinematics and the long-wavelength geoid, our models will provide more realistic constraints on materials properties and slab dynamics. Weak transform faults enhance the motion along the faults. Weak transform faults and their strong adjacent regions cause transform faults to guide plate motion (Fig. 2e and h), consistent with inferences based on plate kinematics[13]. The guiding influence of transform faults on plate motion and the decoupling of thrust faults result in oblique subduction where the strike of thrust faults is non-perpendicular to the transform faults (Figs 2h and 4b), consistent with observations[12,14]. Non-zero normal velocities on faults indicate migrating plate margins. Ultimately, with the formulation presented here and with multiple plates coupled with the solution of energy equation, we should be able to predict the evolution of plate geometry as accomplished in two dimensions[4]. □

Received 12 January; accepted 6 August 1996.

1. Wilson, J. T. Nature **207**, 343–347 (1965).
2. Morgan, J. W. J. Geophys. Res. **73**, 1959–1982 (1968).
3. Hager, B. H. & O'Connell, R. J. J. Geophys. Res. **84**, 1031–1048 (1979).
4. Christensen, U. R. & Harder, H. Geophys. J. Int. **104**, 213–226 (1991).
5. Hager, B. H. & O'Connell, R. J. J. Geophys. Res. **86**, 4843–4867 (1981).
6. Lachenbruch, A. H. & Sass, J. H. Geophys. Res. Lett. **15**, 981–984 (1988).
7. Zhong, S. & Gurnis, M. J. Geophys. Res. **99**, 15683–15695 (1994).
8. Zhong, S. & Gurnis, M. Science **267**, 838–843 (1995).
9. Ranalli, G. *Rheology of the Earth* 326–336 (Chapman & Hall, London, 1995).
10. Moresi, L. N. & Gurnis, M. *Earth Planet. Sci. Lett.* **138**, 15–28 (1996).
11. Gable, C. W., O'Connell, R. J. & Travis, B. J. J. Geophys. Res. **96**, 8391–8405 (1991).

XP-002110730

A genome-wide search for quantitative trait loci underlying asthma

Susan E. Daniels*, Sumit Bhattacharrya*, Alan James†, Nicholas I. Leaves*, Alan Young*‡, Michael R. Hill§, Jennie A. Faux§, Gerrard F. Ryan†, Peter N. le Souef‖, G. Mark Lathrop*, A. William Musk‡ & William O. C. M. Cookson*§

* Wellcome Trust Centre for Human Genetic Disease, University of Oxford, Windmill Road, Oxford OX3 7BN, UK
† Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Perth, Western Australia, Australia
‡ CTSU, Harkness Building, Radcliffe Infirmary, Oxford OX2 6HE, UK
§ Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK
‖ University Department of Paediatrics, Princess Margaret Hospital, Perth, Western Australia, Australia

ASTHMA now affects one child in seven in the United Kingdom[1]. Most cases (95%) of childhood asthma are associated with atopy, the immunoglobulin E (IgE)-mediated familial syndrome of allergic asthma, eczema and rhinitis. Segregation analysis has consistently suggested the presence of major genes influencing atopy and IgE levels[2–4], with the expectation that these genes may be identified by positional cloning or the examination of candidate genes. Here we report the results of a genome-wide search for linkage to one qualitative and four quantitative traits associated with allergic (atopic) asthma. We have identified six potential linkages ($P < 0.001$), five of which are to quantitative traits. Monte Carlo simulations show that 1.6 false-positive linkages at this level of significance would be expected from the data. One linkage, to chromosome 11q13, has been established previously[5]. Three of the new loci show evidence of linkage to a second panel of families, in which maternal effects and pleiotropy of linked phenotypes are seen. The results demonstrate the extent and the complexity of the genetic predisposition to asthma.

Asthma is usually recognized epidemiologically by standard symptom questionnaires or by physician diagnosis[6]. The non-specific bronchial hyper-responsiveness that accompanies

asthma is measured by the dose response of airflow to a bronchoconstrictor such as histamine or methacholine. Atopy is detected by skin prick tests, by measurement of specific serum IgE titres against allergens using radioallergosorbent test (RAST) or enzyme-linked immunosorbent assay (ELISA) techniques, or by quantifying the total serum IgE. Atopic disease is accompanied by blood eosinophilia, and eosinophils are prominent in asthmatic airways.

The prevalence of atopy in Western populations is between 40 and 50%[7-9]. The recruitment of families more than one asthmatic member results in samples in which 70% or more subjects are atopic, making it much more difficult to detect linkage[10]. For this reason, the 80 families in the genome screen (from Busselton in Western Australia) were subselected from a population sample to include non-atopic members. We also reasoned that detection of linkage to quantitative traits would be enhanced by including subjects in normal and abnormal ranges of the trait distributions.

We looked for linkage to four quantitative parameters: the total serum IgE, the skin test index (STI), the peripheral blood eosinophil count, and bronchial responsiveness to methacholine (slope)[11]. A RAST index was also calculated, but gave no additional information to the STI and so was not included in the analysis described below.

Differing indices of atopy may be increased in the same family[12,13], and RAST and skin-test responses reach a peak later in childhood than the total serum IgE, and decline at a slower rate thereafter[9]. To account for this heterogeneity of phenotype (pleiotropy), the categorical trait of 'atopy' was used, based on a combination of STI, RAST index, and total serum IgE. Linkage to quantitative traits was tested by the Haseman–Elston sib-pair technique[14], and to atopy by affected sib-pair methods.

The 80 families surveyed contained a total of 203 offspring forming 172 sib-pairs. The mean age of the children was 12.6 ± 1.3 (s.e.) years, their geometric mean IgE was $55.7 \pm 1.1\ kU\,l^{-1}$, and their mean STI was 4.0 ± 0.41 mm; 12% of the children were asthmatic, and 25% admitted to wheeze; 52% were atopic, 36% were non-atopic, and 12% had an intermediate phenotype as defined; the latter were classified as unknown. The sib-pairs subdivided into 44 pairs in which both siblings were atopic, 87 pairs in which one sibling was affected, and 41 in which both were unaffected or of intermediate phenotype.

The families were initially screened with 269 markers (253 autosomal and 16 X-linked). A genome map was constructed from these data, and the order and distances between the markers were compatible with previously published values[15,16]. Based on these families, the markers had an average heterozygosity of 75%, and spanned approximately 3,100 centimorgans (cM). After initial statistical analysis, none of the X chromosome markers showed evidence of linkage to the five phenotype traits, and the remainder of this analysis therefore deals exclusively with autosomal markers.

TABLE 1 Positive results of genome search (empirical P values)

Marker	θ	$\log_e$IgE	Skin test index	$\log_e$ eosinophil count	$\log_e$ slope	Atopy
D4S171	0.01				$P < 0.05$	
D4S1540	0.10				$P < 0.05$	
D4S426					$P < 0.0005$	
D6S260	0.21			$P < 0.05$		
D6S276	0.03	$P < 0.05$		$P < 0.0001$		
TNFA	0.03					
D6S273	0.06			$P < 0.05$		
D6S291	0.10					$P < 0.005$
D6S271						$P < 0.05$
D7S484	0.00			$P < 0.05$	$P < 0.0005$	
D7S2250	0.03	$P < 0.005$		$P < 0.05$	$P < 0.05$	
D7S528					$P < 0.05$	
FCERB	0.10	$P < 0.0005$	$P < 0.00005$			
FGF3	0.05					
D11S96	0.09	$P < 0.005$				
D11S901		$P < 0.01$				
D13S262	0.04					$P < 0.05$
D13S270	0.04					$P < 0.005$
D13S153	0.20					$P < 0.001$
D13S170						$P < 0.05$
D16S421	0.11	$P < 0.05$				
D16S515	0.12	$P < 0.05$			$P < 0.05$	
D16S516	0.00	$P < 0.05$				
D16S289	0.06	$P < 0.0005$			$P < 0.05$	
D16S507	0.03	$P < 0.005$			$P < 0.05$	
D16S505		$P < 0.05$				

Only regions showing linkage to one or more phenotypes with $P < 0.001$ are shown. The recombination fraction (θ) is to the subsequent marker.

somes 4, 6, 7, 11, 13 and 16, respectively (Table 1). For the regions on chromosomes 4, 7 and 16, the test statistic met the more stringent criterion of $P < 0.0005$, and for two regions, on chromosomes 6 and 11, P was <0.0001. Removal of sib-pair differences that were greater than 3 standard deviations from the mean did not significantly alter the P values for these loci.

The preliminary identification of regions of potential linkage used correlated phenotypes with markers spanning most of the genome. Because of multiple testing, the true significance values will be greater than the normative values given in Table 1. To judge the global significance of the results, a Monte Carlo procedure was applied to calculate the number of regions that would be expected to reach these levels of significance in the absence of any linkages (that is, the number of false-positive linkages). The procedure has advantages over traditional approaches to correcting nominal significance levels because it allows for the informativeness and distance between markers that were actually characterized in the genome-wide search. The expected number of regions of false-positive linkage was 1.6 for $P < 0.001$ (compared with 6 observed linkages), 0.8 for $P < 0.0005$ (5 observed linkages) and 0.2 for $P < 0.0001$ (2 observed linkages), so the observed number of regions meeting these statistical criteria for linkage was much larger than would be expected by chance. In 35,000 simulations, $P < 0.001$ was observed in less than 2% of replicates, indicating that the probability of seeing six or more such regions without a true linkage was <0.005.

Additional markers were studied on the six chromosomes that showed evidence for linkage (Table 1). Although the evidence in favour of linkage was not increased, in all instances at least two of the markers from each region had significance of $P < 0.05$.

the β-chain of the high-affinity receptor for IgE (FcεRI-β), also showed linkage to the STI, but this was not observed on chromosome 16. The regions on chromosome 4 and 7 were linked to bronchial responsiveness, whereas the region on chromosome 6 (near the class I genes of the major histocompatibility complex (MHC)) was linked to eosinophil counts. Weaker evidence of linkage ($P < 0.01$) was found with other phenotypes for the markers on chromosomes 6 and 7 (Table 1).

Markers from chromosome 13, around D13S153, showed evidence of linkage to the atopy phenotype in affected sib-pair analyses (Table 1). The sharing of maternal and paternal alleles identified by descent (IBD) was 46, 1 and 20, 0.

The markers showing $P < 0.001$ for linkage were tested for replication in an additional group (panel B) of 77 nuclear and extended families, recruited from clinics in the United Kingdom, who had previously been used to map atopy on chromosome 11q13[5]. These families had been characterized by the same protocols as the Busselton families, except that bronchial responsiveness and eosinophil counts had not been measured. They contained 215 offspring (268 sib-pairs), of which 61% were atopic and 56% asthmatic, the high values reflecting their origin through clinics. There were 121 atopic-affected sib-pairs, and 70 asthmatic sib-pairs.

Linkage of asthma to FCERB ($P = 0.003$) and to D16S289 ($P = 0.03$) was seen in the panel B families (Table 2). Linkage with atopy was found to D13S153 ($P = 0.003$). The sib-pair sharing of D13S153 alleles in these subjects was 65 IBD = 1, 36 IBD = 0 ($P = 0.003$). The sharing for the combined data was 111 IBD = 1, 56 IBD = 0. The combined Hodge-weighted statistic was 4.158 ($P = 0.000016$). The empirical P value for linkage to D13S153, calculated after 500,000 estimations, was <0.00001.

Many studies have shown that maternal atopy and asthma are more likely to be transmitted to children than is paternal atopy[17-20]. We have not tested for maternal linkage in the genome screen families because of the difficulties induced by a further set of multiple comparisons. However, maternal linkage of atopy to chromosome 11q markers has previously been demonstrated in the panel B families[5,10], and so potentially linked markers identified in the genome screen were tested for parent of origin effects (Table 2).

D4S426, FCERB and D16S289 showed significant differences in linkage between maternal and paternal alleles (Table 2). Linkage to maternal meioses was seen between D4S426 and atopy ($P < 0.05$) and the total serum IgE ($P < 0.001$), and between D16S289 and atopy ($P < 0.01$) and asthma ($P < 0.001$). As already observed, FCERB showed strong maternal linkage to atopy[5,10], but a strong linkage to asthma was also seen ($P < 0.00001$) which has not been previously noted. The finding of maternal effects at several loci favours immunological interactions between mother and child, rather than genetic imprinting or anticipation, as a cause of these phenomena.

It is not clear if linkage of different traits to different regions of the genome indicates that particular genes affect specific traits. However, a previously established region of genetic linkage, to chromosome 11q13, may be used as an examplar[13,21-24]. In the genome screen, the 11q13 marker FCERB was linked to the IgE and the STI. In other studies, FCERB or neighbouring markers have been linked to the composite phenotype of atopy[13,21,22] and to bronchial responsiveness apparently independently of atopy[24]. Similarly, D4S426 and D16S289 showed linkage to different phenotypes in genome screen and the panel B families. Although D13S153 showed linkage to the same phenotype of atopy in both Australian and British families, linkage between the total serum IgE and a polymorphism in the esterase D protein, near D13S153, has previously been suggested, with a lod score of 2.07 (ref. 25).

TABLE 2 Testing for linkage in panel B families: combined and maternal meioses

Locus	IgE	Skin test index	Atopy	Asthma
D4S426				
	($P < 0.001$)*		($P < 0.05$)*	
D6S276				
D7S484				
FCERB				$P < 0.005$
			($P < 0,0001$)**	($P < 0.00001$)*
D13S153			$P < 0.005$	
D16S289				$P < 0.05$
			($P < 0.01$)*	($P < 0.001$)*

Maternal meioses are given in parentheses. Statistical significance values are as follows: * $P < 0.01$ paternal/maternal differences; ** $P < 0.001$ paternal/maternal differences.

tropy may be attributable to environmental events, such as exposure to allergens, and to the age and selection of subjects. The relative distribution of quantitative traits in normal and abnormal ranges may also have differentially influenced the ability to detect linkage.

The cytokine gene cluster on chromosome 5q31 has previously been reported to be linked to serum IgE in Amish families[26], and the result confirmed in Dutch asthmatic families[27]. In our study we did not see linkage to either the serum IgE or to bronchial responsiveness with D5S393, D5S210 or D5S410 from the cytokine gene cluster, although we are currently typing multiple markers within this 20-cM region.

In summary, both the Monte Carlo simulations and the replication of some positive results in a second set of subjects suggest that we have identified regions of true linkage. Replication of linkage to complex traits is to be expected to be difficult[28], and testing for replication in other data sets and with a wider range of markers is still required for the linkages to chromosomes 6 and 7. It is unlikely that all important loci predisposing to atopy and asthma have been mapped, and larger studies and the pooling of data from several centres are required. □

Methods

Subjects. The primary genome screen involved 364 subjects in 80 nuclear families subselected from a population sample of 230 families from Busselton in Western Australia[4,29]. Clinical testing took place over three months in winter, to control for pollen exposure. Young families were recruited to reduce the effects of age and smoking on the IgE levels and lung function. Families in the genome screen included both atopic and non-atopic members, and sibships of three or greater were not exclusively atopic or non-atopic.

Phenotypes. Skin tests to house-dust mite (HDM) and mixed grass pollen (minus the response of negative controls), specific IgE titres to HDM and timothy grass, and the total serum IgE were measured[29]. A 'skin test index' (STI) was calculated as the sum of the skin-prick test results to HDM and grass mix; 95% of individuals in this population who were atopic reacted to either HDM or grass pollen, or both). Bronchial responsiveness to (up to 12 μmol) methacholine was measured[29]. The slope of the dose–response curve was calculated: (pre-dose forced expiratory volume in one second (FEV1) – last FEV1)/cumulative dose of methacholine. A constant of 0.01 was added to each measurement, to allow $\log_e$ transformation when the slope was ≤0. Eosinophils in peripheral blood were Coulter-counted and the values $\log_e$-transformed before analysis. Atopy was defined as STI > 5 mm, or a RAST score to HDM and timothy grass >2, or a total serum IgE > the 7th decile of the age-corrected population. 'Normal' was defined as an STI of 0 and a RAST index of 0, and a total serum IgE < the 7th decile of the age- and sex-matched population. Intermediate phenotypes were classified as unknown. The subjects were administered a modified British MRC questionnaire[29]. Asthma was defined as a positive answer to the questions "Have you ever had an attack of asthma?" and "If yes, has this happened on more than one occasion?"

Genotypes. Genomic DNA was isolated from peripheral blood leucocytes by phenol/chloroform extraction. We typed 22 dinculeotide repeat markers by radioactive techniques[15] (marker names italicized in Table 1). For 274 markers, 50 ng of each DNA sample was genotyped by fluorescence-based semi-automatic methods[15]. Allele sizes were determined using the GENOTYPER soft-

than exact alleles were determined. Non-mendelian marker data were flagged by the programs GAS and UNKNOWN. Potentially incorrect genotypes were re-examined and retested if necessary. Two-point lod scores between markers and marker order were compared with published maps[15,16].

Statistical analysis. Sib-pairs were included in the linkage analysis only when marker genotypes were known for both parents. Linkage was sought to atopy by affected sib-pair analysis. Sibships containing more than one pair were weighted[30]. Linkage to quantitative traits was calculated by the method of Haseman and Elston[14], in which the probabilities that sib-pairs have two, one or no alleles identical by descent are regressed o the square of the sib-pair trait differences[14]. Linkage was examined further in the region of markers where P was <0.001. For Table 1, linkage was presented with $P < 0.01$ or $P < 0.05$ if a marker is <15 cM from a marker with a $P < 0.001$. The P values for linkage with the quantitative trait loci were assessed from the regression analysis of variance (ANOVA), which assumes independence of the sib-pairs. In regions of potential linkage ($P < 0.001$), the P values were validated by empirical calculations in which this assumption is not required. The empirical statistics shown in Table 1 were obtained by simulating the segregation of markers to offspring conditional on the observed parental genotypes without linkage to any of the traits. The empirical P values are the proportion of instances in which the linkage statistic equalled or exceeded its observed value in 500,000 simulated replicates. The frequency of false-positive linkages was assessed by a Monte Carlo simulation, in which replicate sets of observations were simulated under the absence of linkage with mendelian segregation of the offspring genotypes conditional on the observed parental genotypes. Linkage between markers was incorporated with recombination distances equal to those estimated from the data map obtained from this study with offspring simulated under the assumption of lack of interference. The linkage statistics were calculated for each marker locus with the offspring phenotype values as observed. The number of regions in which the linkage statistic for one or more phenotypes exceeded the assigned critical value was calculated over 35,000 replicates. Markers were assumed to detect different regions of linkage if all of the P values were non-significant over a region of >15 cM between two markers with significant linkage.

Replication. Panel B families were recruited from clinics in the United Kingdom[5]. Each family contained at least one asthmatic identified as a proband. Phenotyping was as for the Busselton families, except that eosinophil counts had not been performed and bronchial responsiveness had not been measured. The families were tested with markers that showed potential linkage ($P < 0.001$) to any quantitative or qualitative trait in the Busselton families. Genotyping and linkage analysis were as for the Busselton families. In addition, loci were tested for the presence of significant differences in linkage between maternally and paternally derived alleles; if a significant difference was found, the significance of linkage to maternal alleles was estimated.

Received 4 June; accepted 13 August 1996.

1. Strachan, D. P., Anderson, H. R., Limb, E. S., O'Neill, A. & Wells, N. *Arch. Dis. Child.* **70**, 174–178 (1994).
2. Gerrard, J. W., Rao, D. C. & Morton, N. E. *Am. J. Hum. Genet.* **30**, 46–58 (1978).
3. Borecki, I., Rao, D. C., Lalouel, J. M., McGue, L. & Gerrard, J. W. *Genet. Epidemiol.* **2**, 327–338 (1985).
4. Dizier, M. H. *et al. Genet. Epidemiol.* **12**, 93–105 (1995).
5. Sandford, A. J. *et al. Lancet* **341**, 332–334 (1993).
6. O'Connor, G. T. & Weiss, S. T. *Am. J. Respir. Crit. Care Med.* **149**, S21–28 (1994).
7. Holford-Strevens, V. *et al. J. Allergy Clin. Immunol.* **73**, 516–522 (1984).
8. Peat, J. K., Britton, W. J., Salome, C. M. & Woolcock, A. J. *Clin. Allergy* **17**, 271–281 (1987).
9. Cline, M. G. & Burrows, B. B. *Thorax* **44**, 425–432 (1989).
10. Moffatt, M. F. *et al. Clin. Exp. Allergy* **22**, 1046–1051 (1992).
11. O'Connor, G., Sparrow, D., Taylor, D., Segal, M. & Weiss, S. *Am. Rev. Respir. Dis.* **136**, 1412–1417 (1987).
12. Cookson, W. O. C. M. & Hopkin, J. M. *Lancet* (i), 86–88 (1988).
13. Young, R. P. *et al. J. Med. Genet.* **29**, 236–238 (1992).
14. Haseman, J. K. & Elston, R. C. *Behav. Genet.* **2**, 3–19 (1972).
15. Reed, P. W. *et al. Nature Genet* **7**, 390–395 (1994).
16. Weissenbach, J. *et al. Nature* **359**, 794–801 (1992).
17. Magnusson, C. G. *Allergy* **43**, 241–251 (1988).
18. Halonen, M. *et al. Am. Rev. Respir. Dis.* **146**, 866–870 (1992).
19. Åberg, N. *Clin. Exp. Allergy* **23**, 829–834 (1994).
20. Kuehr, J. *et al. Clin. Exp. Allergy* **23**, 600–605 (1993).
21. Cookson, W. O. C. M., Sharp, P. A., Faux, J. A. & Hopkin, J. M. *Lancet* (i), 1292–1295 (1989).
22. Shirakawa, T. *et al. Clin. Genet.* **46**, 125–129 (1994).
23. Collée, J. M. *et al. Lancet* **341**, 936 (1993).
24. Herwerden, L. *et al. Lancet* **346**, 1262–1265 (1995).
25. Eiberg, H., Lind, P., Mohr, J. & Nielsen, L. S. *Cytogenet. Cell. Genet.* **40**, 622 (1985).
26. Marsh, D. G. *et al. Science* **264**, 1152–1155 (1994).
27. Myers, D. A. *et al. Genomics* **23**, 464–470 (1994).
28. Suarez, B. K., Hampe, C. L. & Van Eerdewegh, P. in *Genetic Approaches to Mental Disorders* (eds Gershon, E. S. & Cloninger, C. R.) 23–46 (American Psychiatric, Washington DC, 1994).
29. Hill, M. R. *et al. BMJ* **311**, 776–779 (1995).
30. Hodge, S. E. *Genet. Epidemiol.* **1**, 109–122 (1984).

ACKNOWLEDGEMENTS. S.E.D. and S.B. contributed equally to this work. We thank the Busselton families and all those who helped to study them, and the Medical Research Council for providing fluorescence-based microsatellite primer sets; non-fluorescent genotyping was in collaboration with INSERM U358. This work was supported by the Wellcome Trust. Additional funding came from the National Asthma Campaign. G.M.L. holds a Wellcome Trust principal fellowship, and W.O.C.M.C. is a Wellcome senior clinical research fellow.

Positional cloning of a global regulator of anterior–posterior patterning in mice

Armin Schumacher, Cynthia Faust & Terry Magnuson

Department of Genetics, Case Western Reserve University, Cleveland, Ohio 44106, USA

ANTERIOR-POSTERIOR (A–P) patterning is of fundamental importance throughout vertebrate embryonic development. Murine members of the *trithorax* (*trx*) and *Polycomb* group (*Pc*-G) of genes regulate A–P patterning of segmented axial structures[1–4], demonstrating conserved upstream regulation of homeotic pathways between *Drosophila* and mouse. Here we report the positional cloning of a classical mouse mutation, *eed* (for *embryonic ectoderm development*), which is the highly conserved homologue of the *Drosophila Pc*-G gene *esc* (for *extra sex combs*), a long-term repressor of homeotic genes[5]. Mutants homozygous for a null allele of *eed* display disrupted A–P patterning of the primitive streak during gastrulation. Mutant embryos lack a node, notochord and somites, and there is no neural induction[6]. In contrast to absence of anterior structures, extra-embryonic mesoderm is abundant. Mice carrying a hypomorphic *eed* mutation exhibit posterior transformations along the axial skeleton. These results indicate that *eed* is required globally for A–P patterning throughout embryogenesis.

The recessive, embryonic-lethal mutation *eed* maps to the albino-deletion complex on chromosome 7 (ref. 7). Previous genetic and molecular analyses positioned *eed* within a 150-kilobase (kb) interval between the c^{3H} and the c^{4L} deletion breakpoints[7] (Fig. 1*a*). Here we have isolated *eed* from a CpG island (that is, a cluster of methylation-sensitive rare-cutter sites) located approximately 40 kb proximal to the c^{4L} deletion breakpoint[7]. Molecular access was obtained by chromosomal walking from the proximal end of yeast artificial chromosome (YAC) D6E7 (Fig. 1*a*). A 1.1-kb *SalI* fragment from the CpG island was used to isolate overlapping clones from a mouse embryonic day 7.5 (E7.5) cDNA library generating a 1.95-kb cDNA (Fig. 1*a*). The candidate cDNA was validated as the *eed* gene by analysing two non-complementing ethyl-*N*-nitrosurea (ENU)-induced alleles ($l7Rn5^{3354SB}$ and $l7Rn5^{1989SB}$) at the *eed* locus[8,9]. The $l7Rn5^{3354SB}$ and $l7Rn5^{1989SB}$ mutations represent null[10] and hypomorphic alleles[9], respectively. DNA sequence analysis of reverse transcription–polymerase chain reaction (RT–PCR) products from homozygous E8.5 $l7Rn5^{3354SB}$ embryos showed a transition of thymine at position 1040 to cytosine ($T^{1040} \rightarrow C$) (Fig. 1*a*). This base exchange was absent in an ENU-induced allele, $l7Rn5^{4234SB}$, which *trans*-complements the $l7Rn5^{3354SB}$ allele[9]. Polymorphism can be excluded because both mutations were generated on the same BALB/cR1 background[8]. As a second co-isogenic control, RT–PCR products from E12.5 homozygous $l7Rn5^{1989SB}$ embryos also showed the expected wild-type T^{1040}. The sequence alteration in $l7Rn5^{1989SB}$ was identified as a $T^{1031} \rightarrow A$ transversion which was absent in both $l7Rn5^{4234SB}$ and $l7Rn5^{3354SB}$ (Fig. 1*a*).

Conceptual translation revealed an open reading frame of 441 amino acids, containing five tandemly repeated WD motifs (Fig. 1*a*). A potential PEST sequence[11] is defined by residues 22 to 52 of Eed protein (Fig. 1*a*). Two discrete clusters of basic amino acids encompassing a nine-residue spacer region (RK-X_9-KKXKX, residues 65 to 80; Fig. 1*a*) are reminiscent of a bipartite nuclear targeting signal[12]. Eed does not display any canonical DNA-binding motif. Three putative Asn-linked glycosylation sites (NXS/T)[13] at amino-acid positions 157, 283 and 349 (Fig. 1*a*) are compatible with a potential non-nuclear isoform of Eed